Evaluation of CRU TS, GPCC, AgMERRA, and AgCFSR meteorological datasets for estimating climate and crop variables: A case study of maize in Qazvin Province, Iran

Faraz GORGIN PAVEH[1*], Hadi RAMEZANI ETEDALI[2], Brian COLLINS[3]

[1] Syracuse University, Syracuse 13244, USA;
[2] Imam Khomeini International University, Qazvin 34149-16818, Iran;
[3] Centre for Crop Science, The University of Queensland, Brisbane 4072, Australia

Abstract: In the past few decades, meteorological datasets from remote sensing techniques in agricultural and water resources management have been used by various researchers and managers. Based on the literature, meteorological datasets are not more accurate than synoptic stations, but their various advantages, such as spatial coverage, time coverage, accessibility, and free use, have made these techniques superior, and sometimes we can use them instead of synoptic stations. In this study, we used four meteorological datasets, including Climatic Research Unit gridded Time Series (CRU TS), Global Precipitation Climatology Centre (GPCC), Agricultural National Aeronautics and Space Administration Modern-Era Retrospective Analysis for Research and Applications (AgMERRA), Agricultural Climate Forecast System Reanalysis (AgCFSR), to estimate climate variables, i.e., precipitation, maximum temperature, and minimum temperature, and crop variables, i.e., reference evapotranspiration, irrigation requirement, biomass, and yield of maize, in Qazvin Province of Iran during 1980–2009. At first, data were gathered from the four meteorological datasets and synoptic station in this province, and climate variables were calculated. Then, after using the AquaCrop model to calculate the crop variables, we compared the results of the synoptic station and meteorological datasets. All the four meteorological datasets showed strong performance for estimating climate variables. AgMERRA and AgCFSR had more accurate estimations for precipitation and maximum temperature. However, their normalized root mean square error was inferior to CRU for minimum temperature. Furthermore, they were all very efficient for estimating the biomass and yield of maize in this province. For reference evapotranspiration and irrigation requirement CRU TS and GPCC were the most efficient rather than AgMERRA and AgCFSR. But for the estimation of biomass and yield, all the four meteorological datasets were reliable. To sum up, GPCC and AgCFSR were the two best datasets in this study. This study suggests the use of meteorological datasets in water resource management and agricultural management to monitor past changes and estimate recent trends.

Keywords: climate variables; crop variables; meteorological datasets; precipitation; reference evapotranspiration; irrigation requirement; Iran

Citation: Faraz GORGIN PAVEH, Hadi RAMEZANI ETEDALI, Brian COLLINS. 2022. Evaluation of CRU TS, GPCC, AgMERRA, and AgCFSR meteorological datasets for estimating climate and crop variables: A case study of maize in Qazvin Province, Iran. Journal of Arid Land, 14(12): 1361–1376. https://doi.org/10.1007/s40333-022-0108-7

*Corresponding author: Faraz GORGIN PAVEH (E-mail: fgorginp@syr.edu)
Received 2022-07-04; revised 2022-11-02; accepted 2022-11-17

1 Introduction

The impact of climate change on agricultural product yields is undeniable, and changing precipitation and temperature can affect agricultural production and time (Massah and Morid, 2005). Accurate temperature and precipitation estimations and historical change tracking will substantially aid in adapting to threats. On the other hand, accurate estimation of crop variables, such as reference evapotranspiration, irrigation requirement, biomass, and yield, is undeniable. Ground stations currently provide insufficient temporal and geographical data in many parts of the world (Worqlul et al., 2015). Nowadays, numerous options are provided to observe the water conditions around the world, and datasets are one of them (Ramezani Etedali and Ahmadi, 2021).

Models simulating water and crops can estimate agricultural water requirements, yield, and other crop indicators. Although several crop simulation models have been reported in the literature that quantifies the effect of stresses on crop growth, development, and yield, such as Cropping Systems Simulation (CropSyst) model (Stöckle et al., 2003), Agricultural Production Systems sIMulator (APSIM) model (Keating et al., 2003), Hybrid-Maize model (Yang et al., 2004), and a computer program for irrigation planning and management (CROPWAT) (El-mageed et al., 2017). The typical problem of these models is that they require very comprehensive input data and information on crop growth, which may not be available in most parts of the world (Queyrel et al., 2016). In the past recent years, there have been various researches related to the prediction of climate and crop models in Iran as well as other countries, besides the significant outcomes, there are problems of lack of accuracy and uncertainty, which shows the importance of more research in this field (Javanmard et al., 2010; Katiraie-Boroujerdy et al., 2013).

The cost-effective wide range and time-saving initiatives can be done locally, regionally, nationally, and worldwide by developing satellites and using remote sensing techniques (Shi et al., 2017). They also encompass data across various time intervals, from hourly to daily, monthly, and yearly. Moreover, they make it feasible to examine changes and monitor terrestrial events using these data (Ge et al., 2019), including agricultural and natural resources management (Rodriguez et al., 2015), groundwater (Ollivier et al., 2021), water quality (Li et al., 2021), salinity management (Apan et al., 2004), and crop water requirement (Salehnia et al., 2018; Olivera Rodriguez et al., 2021).

Because of the need to provide global scale hydrological parameters, land surface-updated models are critical. These datasets are not more accurate than the classic measurements, but they have advantages in several aspects, the most important of which is full spatial and temporal coverage. In terms of application, this knowledge is separated into three categories. The first category includes data from previous years, as this information is useful in understanding how the water cycle evolves through time. The study of rainfall and drought trends is another application of these datasets. The data in these databases enable continuous evaluation of the changing water resource process and, if necessary, management decisions (Bowen et al., 2019; Vorosmarty and Sahagian, 2000). Short-term data for meteorological measurements are the second category (Benevides et al., 2019; Kreuzer et al., 2020). This data is also crucial for making decisions in short periods. The third category is offered through the application of models in different circumstances, which provides information about future precipitation and can be used for long-term management (Worqlul et al., 2015; Yao et al., 2019).

For agricultural management, precise data in numerous fields, such as crop yield, are required. Estimation of yield is now regarded as an essential aspect of agriculture. Climate data are a crucial component of plant model input data. Different databases of various types might help compensate for these flaws. However, the accuracy of the datasets must be assessed. In this study, we estimated minimum temperature, maximum temperature, and precipitation in Qazvin Province of Iran during 1980–2009 by four meteorological datasets, including Climatic Research Unit gridded Time Series (CRU TS), Global Precipitation Climatology Centre (GPCC), Agricultural National Aeronautics and Space Administration Modern-Era Retrospective Analysis for Research and Applications (AgMERRA), and Agricultural Climate Forecast System Reanalysis (AgCFSR).

The results are compared to the synoptic station in Qazvin Province by evaluation indices, such as correlation coefficient (R^2), root mean square error (RMSE), normalized root mean square error (NRMSE), and maximum error (ME). Then, the four meteorological datasets are used to estimate the reference evapotranspiration, biomass, irrigation requirement, and yield of maize in Qazvin Province on a monthly scale. Finally, the results are compared to the data from synoptic station in Qazvin Province with the mentioned indices.

2 Study area and methods

2.1 Qazvin Province and maize

Qazvin Province is located in northwestern Iran, south of the Alborz Mountains, with high latitudes. Although the province covers only 1% of the country's area (1.6×10^4 km^2), it provides about 3% of Iran's agricultural production, with more than 5.9×10^6 t (Ordikhani et al., 2021). The average annual precipitation of the province is 330 mm (from 210 mm in the east to 550 mm in the northeast). The province includes two parts: the plain zone and the mountain zone. Table 1 provided information about maximum temperature and minimum temperature on a monthly scale in Qazvin Province. In addition, daily evapotranspiration, monthly evapotranspiration, and precipitation are also provided in this table.

Table 1 Information about the climate variables in Qazvin Province during study period

Month	Maximum temperature (℃)	Minimum temperature (℃)	Daily evapotranspiration (mm)	Monthly evapotranspiration (mm)	Precipitation (mm)
January	6.2	–4.1	0.9	29.3	35.8
February	8.6	–2.5	1.6	44.2	40.4
March	14.1	1.4	2.7	84.7	51.1
April	20.6	6.6	4.0	119.5	47.4
May	26.0	10.3	5.3	163.6	30.9
June	32.6	14.8	7.5	224.6	4.2
July	35.4	17.6	7.8	242.7	3.3
August	34.9	17.1	7.2	221.8	8.7
September	30.8	13.2	5.5	165.8	1.2
October	23.4	8.3	3.3	100.8	28.1
November	14.6	3.0	1.6	47.6	44.7
December	8.3	–1.8	0.9	28.3	43.6

Wheat, barley, maize, and alfalfa are principal crops in Qazvin Province, which are valuable in Iran. Information from the synoptic station of the province (36°8′59″N, 50°1′47″E; 1279.2 m) was gathered in this study (Fig. 1). In this study, we collected monthly and annually data from 1980 to 2009. As maize covers more than 1.3×10^2 km^2 of agricultural land in Qazvin Province, it can represent the province's agriculture crop well. Therefore, in this study, we used maize as a sample of the province. In addition, we used data from the synoptic station instead of the actual documented data on maize in the province. The reason is that the synoptic station provides data for the long-term period. However, in order to obtain reliable data, we calibrated the synoptic station data based on the actual data (Ramezani Etedali and Ahmadi, 2021; Saeidi et al., 2021).

2.2 AquaCrop model

In this study, we used AquaCrop model to calculate the reference evapotranspiration, biomass, irrigation requirement, and crop yield of maize from the synoptic station and meteorological datasets.

Food and Agriculture Organization of the United Nations (FAO) developed AquaCrop model, a crop development model, to examine the crop's environmental impacts and water management.

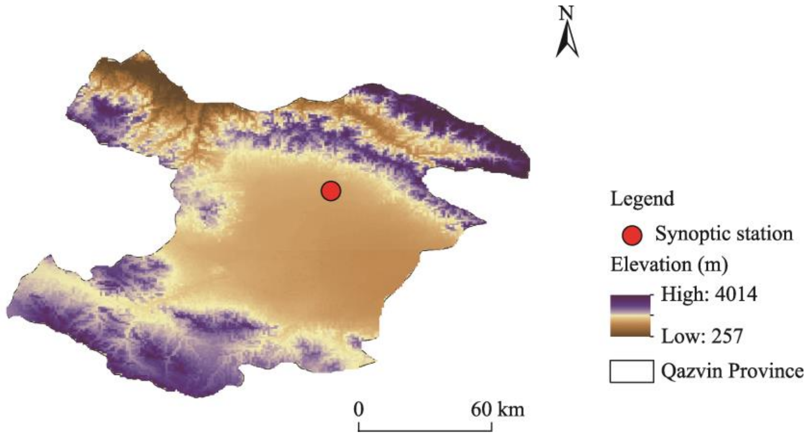

Fig. 1 Study area and the location of synoptic station used in the study

AquaCrop model is a water-driven model that affects crop yields by regulating the amount of available water in the soil. AquaCrop model also has the advantage of requiring fewer input variables than other models and maintains the required precision (Confalonieri et al., 2016). AquaCrop model has been proven in numerous experiments to accurately simulate crop production under various water management circumstances (Abedinpour et al., 2012; Hellal et al., 2019). The empirical equation of AquaCrop model is as follows (Dorenbos and Kassam, 1979):

$$\left(1-\frac{Y}{Y_x}\right)=K_y\left(1-\frac{ET}{ET_x}\right), \tag{1}$$

where Y_x and Y represent maximum and actual crop production (kg/m^2), respectively; ET_x and ET represent potential and actual evapotranspiration (mm), respectively; and K_y represents the conversion coefficient between crop yield and water stress.

To avoid any confusion, the model separates evapotranspiration into soil evaporation and crop transpiration (Zhu et al., 2021). Reference evapotranspiration (ET_0) was calculated based on the FAO's Penman-Monteith Equation (Allen et al., 1998):

$$ET_0=\frac{0.408\Delta(R_n-G)+\gamma(900/(T+273))u_2(e_s-e_a)}{\Delta+\gamma(1+0.34u_2)}, \tag{2}$$

where R_n is the net radiation at the surface of the crop (MJ/(m^2•a)); G is the soil heat flux density (MJ/(m^2•a)); T is the air temperature at a height of two meters (℃); u_2 is the wind speed at a height of two meters (m/s); e_s is the saturation vapour pressure of the air (kPa); e_a is the actual vapour pressure (kPa); Δ is the slope of vapour pressure curve (kPa/℃); and γ is the psychrometric constant (kPa/℃).

AquaCrop model requires four types of information as input data: (1) meteorological data, comprising daily precipitation, maximum temperature, minimum temperature, and reference evapotranspiration from FAO's Penman-Monteith Equation (Allen et al., 1998); (2) soil data, including the texture and organic content of the soil; (3) crop data, involving the factors of growth and development, evaporation, transpiration, yield, and stress; and (4) information on management, containing irrigation, fertilizer, and surface covering (Golabi and Naseri, 2015; Zhu et al., 2021). In this study, the soil texture of Qazvin Province is considered as loam, with 32.2% field capacity and 16.1% permanent wilting point.

2.3 Meteorological datasets

2.3.1 Climatic Research Unit gridded Time Series (CRU TS)

CRU TS contains 10 observed and calculated variables and provides a monthly grid of land-based observations dating back to 1901. In the defined domain, there are no missing values.

Precipitation, maximum temperature, and minimum temperature are some of the accessible data from this dataset (New et al., 1999; Harris et al., 2020). CRU TS was initially released in 2000, and used Angular Distance Weighting Interpolation (ADW) to interpolate monthly anomalies for seven variables onto a 0.50° grid. After thoroughly reviewing the possibilities, ADW was chosen as the interpolation method. The observed coverage was 'plugged' by interpolating the synthesized data onto a coarser grid (2.50°) (New et al., 2000; Mitchell et al., 2004; Mitchell and Jones, 2005; Harris et al., 2014).

2.3.2 Global Precipitation Climatology Centre (GPCC)

GPCC was created in 1989 as the in situ component of The Global Energy and Water Exchanges (GEWEX)'s Global Precipitation Climatology Project (WMO et al., 1990). Its primary responsibility is to analyze monthly precipitation for the Earth's land surface using rain gauge (in situ) observations. It collected a unique database of precipitation data from more than 85,000 stations around the world that is estimated to be between 150,000 and 250,000 rain gauges by previous studies (New et al., 2001; Strangeways, 2006; Schneider et al., 2014).

2.3.3 Agricultural National Aeronautics and Space Administration Modern-Era Restrospective Analysis for Research and Applications (AgMERRA)

AgMERRA was created to provide daily time series on a global basis during 1980–2009 (Ruane et al., 2015). At spatial resolution of 0.25°, AgMERRA provides daily precipitation, maximum temperature, minimum temperature, relative humidity, and solar radiation. Reichle et al. (2011) and Ruane et al. (2015) provided more information about AgMERRA.

2.3.4 Agricultural Climate Forecast System Reanalysis (AgCFSR)

AgCFSR includes rainfall data with daily time and 0.25° geographic resolution during 1980–2009. CRU TS, GPCC, Tropical Rainfall Measuring Mission (TRMM), and Climate Prediction Center Morphing technique (CMORPH) databases are used to create AgCFSR, which combines in-situ and satellite data. It also considers the agricultural areas and climatic elements affecting crops (Ruane et al., 2015).

The information of all the four datasets is summarized in Table 2.

Table 2 Information about the grid resolution and company's name of the four datasets used in this study

Dataset	Grid resolution	Company's name
CRU TS	0.50°	United Kingdom's Natural Environment Research Council and United Kingdom National Centre for Atmospheric Science
GPCC	1.00°	German Weather System
AgMERRA	0.25°	United States National Aeronautics and Space Administration
AgCFSR	0.25°	United States National Aeronautics and Space Administration

Note: CRU TS, Climatic Research Unit gridded Time Series; GPCC, Global Precipitation Climatology Centre; AgMERRA, Agricultural National Aeronautics and Space Administration Modern-Era Retrospective Analysis for Research and Applications; AgCFSR, Agricultural Climate Forecast System Reanalysis.

2.4 Evaluation indices

We need to understand how much data are need to use meteorological datasets to replace synoptic stations. For this purpose, we tried to find the datasets close to the synoptic station of the province to understand the best efficient method to estimate meteorological and crop models. Data from datasets were measured in three methods: (1) only the closest point of the datasets to the synoptic station was measured (K1); (2) the average of the four closest points of the datasets to the synoptic station was calculated (K4); and (3) the average of the eight nearest points of the datasets to the synoptic station was calculated (K8). In other words, data were gathered from the closest point, the four closest points, and the eight nearest points to the synoptic station in the province. We gathered data from the closest point, the four closest points, and the eight closest points to the synoptic station, and then averaged these data before comparing them with the synoptic station data.

For the data of precipitation, maximum temperature, minimum temperature, the K8 method was not helpful as it decreased the efficiency of the estimations. However, it was relatively efficient in increasing the accuracy of the estimations of crop variables. Therefore, we only provided data from K1 and K4 for climate variables and from K1, K4, and K8 for crop variables. In addition, in this study, we used monthly data from datasets to estimate climate variables, then we used the averaged data to estimate crop variables annually.

The maize production and reference evapotranspiration were assessed using CRU TS, GPCC, AgMERRA, and AgCFSR meteorological datasets and compared to the synoptic station estimation. Outcomes were compared using statistical error criteria, such as R^2, RMSE, NRMSE, and ME.

3 Results

3.1 Climate variables

3.1.1 Minimum temperature

Figure 2 provided information about minimum temperature estimation for K1 and K4 in Qazvin Province from 1980 to 2009. As GPCC does not provide information about temperature, there is no information about this dataset for minimum temperature and maximum temperature.

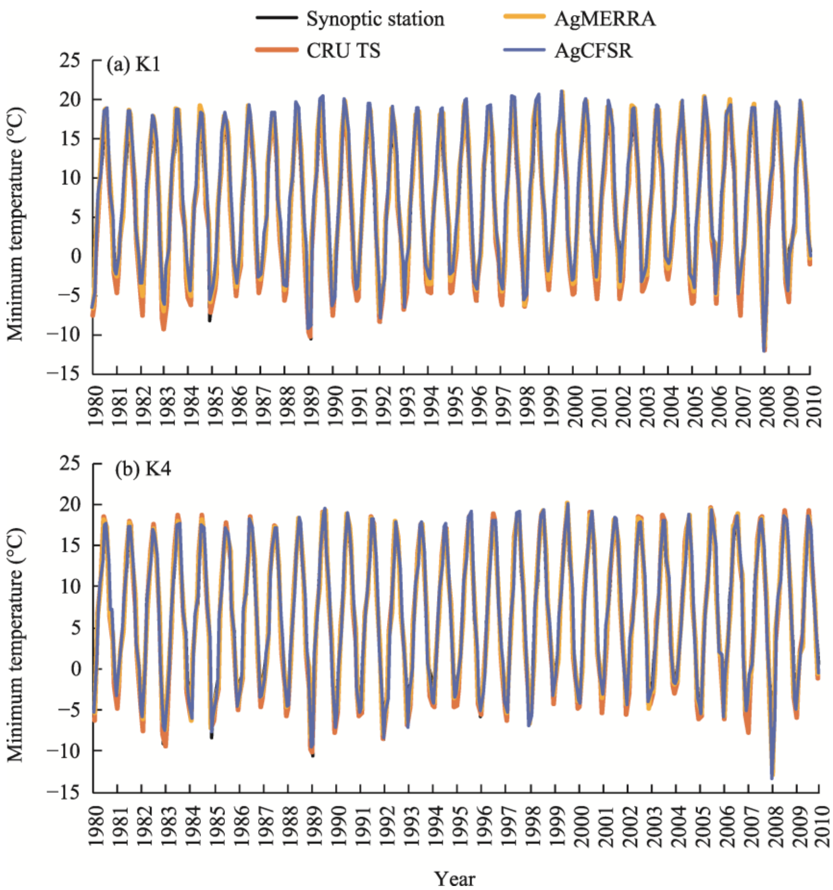

Fig. 2 Comparison of minimum temperature estimated by Climatic Research Unit gridded Time Series (CRU TS), Agricultural National Aeronautics and Space Administration Modern-Era Retrospective Analysis for Research and Applications (AgMERRA), and Agricultural Climate Forecast System Reanalysis (AgCFSR) using K1 (a) and K4 (b) with the data from the synoptic station in Qazvin Province during 1980–2009. K1 represents only the closest point of the datasets to the synoptic station was measured, and K4 represents the average of the four closest points of the datasets to the synoptic station was calculated.

According to the figure, all the three datasets (CRU TS, AgMERRA, and AgCFSR) efficiently estimated minimum temperature; moreover, they estimated close to each other. Statistical evaluations of these datasets were provided in Table 3. As can be seen from this table, all datasets had a solid R^2 for estimating minimum temperature, as they were above 0.99 for both K1 and K4. For K1, CRU TS was the most substantial dataset with an NRMSE of 5.96%, while the other two datasets, i.e., AgMERRA and AgCFSR, had an NRMSE value of 10.55% and 14.51%, respectively. Although the two datasets had a high NRMSE value, their ME was relatively low. On the other hand, the results were different for K4, as they had differences in RMSE and NRMSE. The value of RMSE slightly increased for CRU TS but significantly decreased for the other two datasets. We can observe same behavior for the changes of NRMSE, as it decreased slightly from around 6.00% for K1 to 7.55% for K4 in CRU TS and decreased significantly in about 4.00% in both AgMERRA and AgCFSR.

Table 3 Comparison of minimum temperature, maximum temperature, and precipitation estimated by CRU TS, GPCC, AgMERRA, and AgCFSR with the data from the synoptic station in Qazvin Province during 1980–2009

Dataset	Minimum temperature							
	R^2		RMSE (℃)		NRMSE (%)		ME (℃)	
	K1	K4	K1	K4	K1	K4	K1	K4
CRU TS	0.99	0.99	0.4	0.5	5.96	7.55	2.6	3.0
AgMERRA	0.99	0.99	0.7	0.3	10.55	4.33	3.1	3.2
AgCFSR	0.99	0.99	1.0	0.3	14.51	4.30	1.1	2.1
Dataset	**Maximum temperature**							
	R^2		RMSE (℃)		NRMSE (%)		ME (℃)	
	K1	K4	K1	K4	K1	K4	K1	K4
CRU TS	0.99	0.99	1.3	1.3	6.32	6.62	4.3	4.5
AgMERRA	0.99	0.99	0.2	1.0	1.10	4.80	4.9	6.3
AgCFSR	0.99	0.99	0.8	1.1	3.82	5.09	1.4	3.4
Dataset	**Precipitation**							
	R^2		RMSE (mm)		NRMSE (%)		ME (mm)	
	K1	K4	K1	K4	K1	K4	K1	K4
CRU TS	0.62	0.62	1.9	3.3	6.89	11.90	182.6	182.0
GPCC	0.72	0.70	2.4	3.2	8.53	11.80	192.7	182.6
AgMERRA	0.10	0.60	1.6	3.3	5.61	11.78	167.8	187.4
AgCFSR	0.56	0.66	1.5	3.5	5.43	11.79	187.9	187.3

Note: R^2, correlation coefficient; RMSE, root mean square error; NRMSE, normalized root mean square error; ME, maximum error. K1 represents only the closest point of the datasets to the synoptic station was measured; K4 represents the average of the four closest points of the datasets to the synoptic station was calculated; and K8 represents the average of the eight nearest points of the datasets to the synoptic station was calculated.

3.1.2 Maximum temperature

Figure 3 gived information about maximum temperature estimated by CRU TS, AgMERRA, and AgCFSR datasets in Qazvin Province from 1980 to 2009. According to the figure, all datasets estimated maximum temperature with high accuracy and were close to each other. In addition, there was a good correlation between all datasets and the synoptic station, as the value of R^2 was above 0.99 for K1 and K4. Also, the value of RMSE was low for all datasets and NRMSE showed high accuracy, as it was 6.32% for CRU TS, 3.82% for AgCFSR, and 1.10% for AgMERRA for K1. Although AgMERRA was the best dataset for estimating maximum temperature, the lowest value of ME was AgCFSR. Results were less efficient but still suitable for K4. As a result, the best dataset was AgMERRA, with an NRSME of 4.80%. CRU TS was the least efficient dataset with an NRMSE of 6.32%, second only to AgCFSR with an NRMSE of 5.09%.

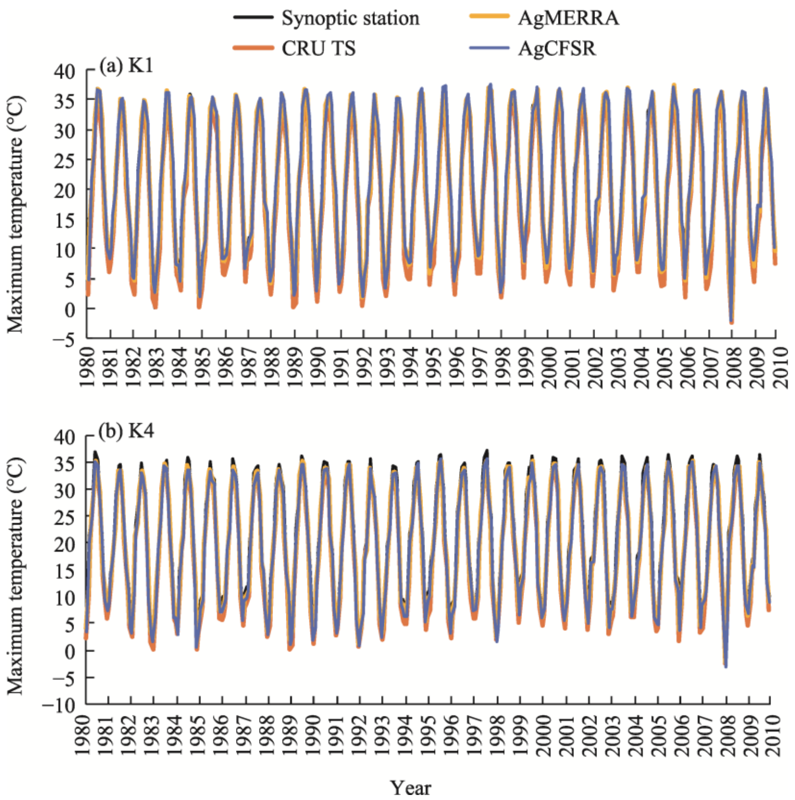

Fig. 3 Comparison of maximum temperature estimated by CRU TS, AgMERRA, and AgCFSR using K1 (a) and K4 (b) with the data from the synoptic station in Qazvin Province during 1980–2009

3.1.3 Precipitation

Figure 4 provided information about precipitation variation in Qazvin Province estimated by the four meteorological datasets and from the synoptic station in the province during 1980–2009. This figure showed that precipitation estimated by datasets strongly correlated to precipitation from the synoptic station for both K1 and K4. Also, statistical indices were given in Table 3 for all the four datasets. According to Table 3, unlike AgMERRA, datasets strongly correlated (above 0.50) to synoptic station for K1, while for K4, R^2 of all datasets was above 0.50. Moreover, although the value of ME was relatively high, the value of NRMSE for all of them was less than 10.00%, with RMSE around 2.0 mm for K1 and around 3.0 mm for K4. The better datasets were AgCFSR and AgMERRA, and CRU TS and GPCC ranked third and fourth place, respectively, and they all showed relatively the same behavior for K4.

According to the above results, we can conclude that all datasets were efficient for estimating climate variables in Qazvin Province. They all showed a high correlation to synoptic station in the province, and the value of NRMSE was efficient enough. However, we can choose AgMERRA and AgCFSR as the most suitable datasets for estimating maximum temperature and precipitaion and CRU TS as the most suitable datasets for estimating minimum temperature.

3.2 Crop variables

3.2.1 Reference evapotranspiration

GPCC did not provide the related data estimating reference evapotranspiration. Therefore, we offered the results of CRU TS, AgMERRA, and AgCFSR for estimating reference evapotranspiration of maize in this study (Fig. 5). The statistical evaluation of datasets estimations was provided in Table 4. CRU TS estimated reference evapotranspiration efficiently with RMSE around 156.00 mm to 172.00 mm for K1, K4, and K8, and an NRMSE value of less than 15.00%. ME values were close to each other, around 360.00 mm; however, R^2 was not high.

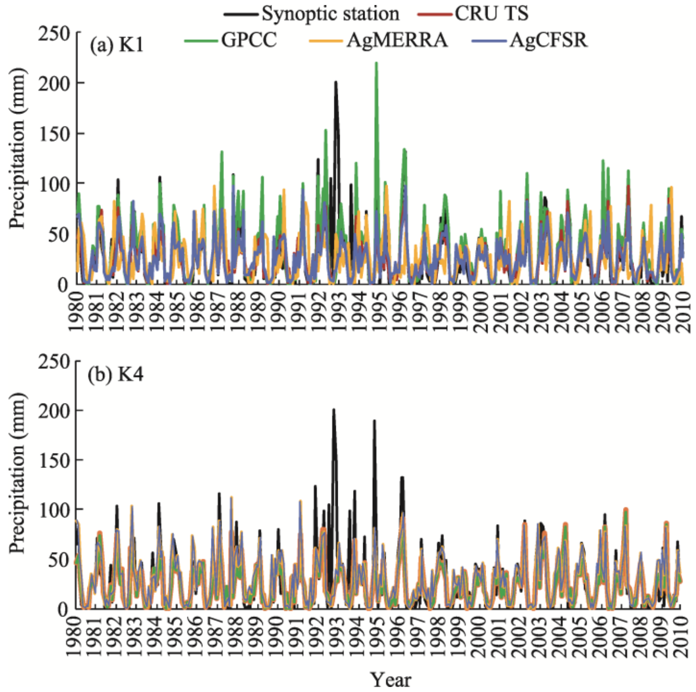

Fig. 4 Comparison of precipitation estimated by CRU TS, Global Precipitation Climatology Centre (GPCC), AgMERRA, and AgCFSR using K1 (a) and K4 (b) with the data from the synoptic station in Qazvin Province during 1980–2009

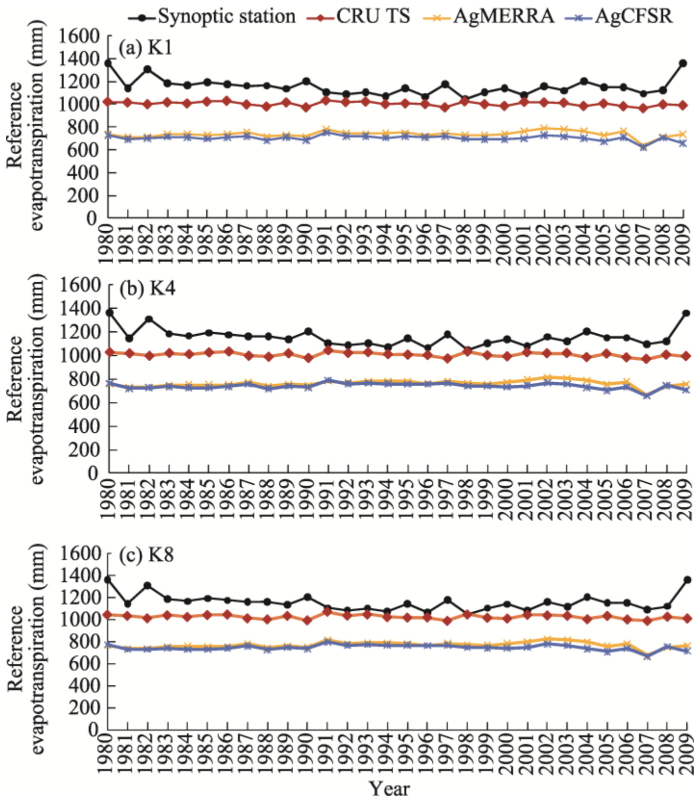

Fig. 5 Comparison of reference evapotranspiration estimated by CRU TS, AgMERRA, and AgCFSR using K1 (a), K4 (b), and K8 (c) with the data from the synoptic station in Qazvin Province during 1980–2009. K8 represents the average of the eight closest points of the datasets to the synoptic station was calculated.

According to Table 4, R^2 was not good at any part of this section, even though other indices were considered efficient and reliable. Therefore, we neglected R^2 in this section and focused on different indices. For CRU TS, it was clear that we can find a more efficient method by choosing more points to estimate reference evapotranspiration (i.e., K8 and K4). The indices for the other two datasets, AgMERRA and AgCFSR, were close to each other, but AgMERRA was slightly better than AgCFSR. Furthermore, K4 made the best estimation for reference evapotranspiration in compare with K1 and K8 in this province.

Table 4 Statistical evaluation of reference evapotranspiration, irrigation requirement, biomass, and yield estimated by CRU TS, GPCC, AgMERRA, and AgCFSR in Qazvin Province from 1980 to 2009

Dataset	Reference evapotranspiration											
	R^2			RMSE (mm)			NRMSE (%)			ME (mm)		
	K1	K4	K8	K1	K4	K8	K1	K4	K8	K1	K4	K8
CRU TS	0.02	0.30	0.30	172.45	168.28	156.31	14.93	14.57	13.24	370.00	347.00	365.00
GPCC	-	-	-	-	-	-	-	-	-	-	-	-
AgMERRA	0.00	0.02	0.02	425.32	403.00	593.30	36.83	34.90	51.38	625.00	595.00	604.00
AgCFSR	0.00	0.03	0.03	460.11	424.00	616.46	39.85	36.72	53.39	701.00	645.00	653.00
Dataset	Irrigation requirement											
	R^2			RMSE (mm)			NRMSE (%)			ME (mm)		
	K1	K4	K8	K1	K4	K8	K1	K4	K8	K1	K4	K8
CRU TS	0.00	0.00	0.01	130.72	127.89	135.04	18.28	17.88	18.88	255.00	281.00	283.00
GPCC	0.08	0.02	0.06	145.00	150.86	144.00	20.27	21.09	20.14	323.00	310.00	329.00
AgMERRA	0.07	0.21	0.11	305.69	310.98	507.39	42.74	43.48	70.94	457.00	526.00	531.00
AgCFSR	0.00	0.01	0.05	311.89	316.05	528.13	43.61	44.19	73.84	255.00	499.00	538.00
Dataset	Biomass											
	R^2			RMSE (t/hm^2)			NRMSE (%)			ME (t/hm^2)		
	K1	K4	K8	K1	K4	K8	K1	K4	K8	K1	K4	K8
CRU TS	0.10	0.15	0.14	1.19	0.90	0.60	4.03	3.02	2.12	2.14	1.74	1.35
GPCC	0.13	0.15	0.15	1.09	0.80	0.60	3.69	2.89	2.04	1.88	1.62	1.29
AgMERRA	0.05	0.06	0.07	1.29	0.90	0.70	4.39	2.90	2.54	2.78	2.13	1.88
AgCFSR	0.04	0.13	0.12	1.15	0.70	0.70	3.89	2.52	2.38	2.62	1.60	2.00
Dataset	Yield											
	R^2			RMSE (t/hm^2)			NRMSE (%)			ME (t/hm^2)		
	K1	K4	K8	K1	K4	K8	K1	K4	K8	K1	K4	K8
CRU TS	0.08	0.08	0.09	0.61	0.44	0.34	4.27	3.13	2.41	1.31	0.99	0.77
GPCC	0.10	0.13	0.09	0.56	0.43	0.33	3.96	3.07	2.34	1.60	0.95	0.73
AgMERRA	0.01	0.01	0.01	0.68	0.46	0.42	4.80	3.26	2.96	1.18	1.10	1.12
AgCFSR	0.06	0.02	0.02	0.59	0.42	0.39	4.14	2.94	2.78	1.26	0.96	1.00

Note: - denotes no data.

3.2.2 Irrigation requirement

Irrigation requirement was calculated to show the difference between evapotranspiration and efficient precipitation. This variable was calculated and estimated to assess all the climate variables from datasets. The results of the estimating irrigation requirement of maize were represented in Figure 6. Statistical evaluation of these estimations was provided in Table 4. According to Figure 6, the results of CRU TS and GPCC were close with high reliability, while the results of AgMERRA and AgCFSR were close with less accuracy. As seen in Table 4, CRU

TS was more efficient for estimating irrigation requirement of maize in Qazvin Province with RMSE of less than 135.00 mm, NRMSE of about 18.00%, and ME of about 270.00 mm. GPCC dataset was slightly less accurate than CRU TS, with NRMSE of about 20.00% and RMSE of less than 150.00 mm. K4 method was more accurate for CRU TS, but for GPCC, there was no significant difference among K1, K4, and K8. The other two datasets, i.e., AgMERRA and AgCFSR, were not as efficient as the first two; their RMSE and NRMSE were about 300.00 mm and 40.00% for K1 and K4 and much less accurate for K8.

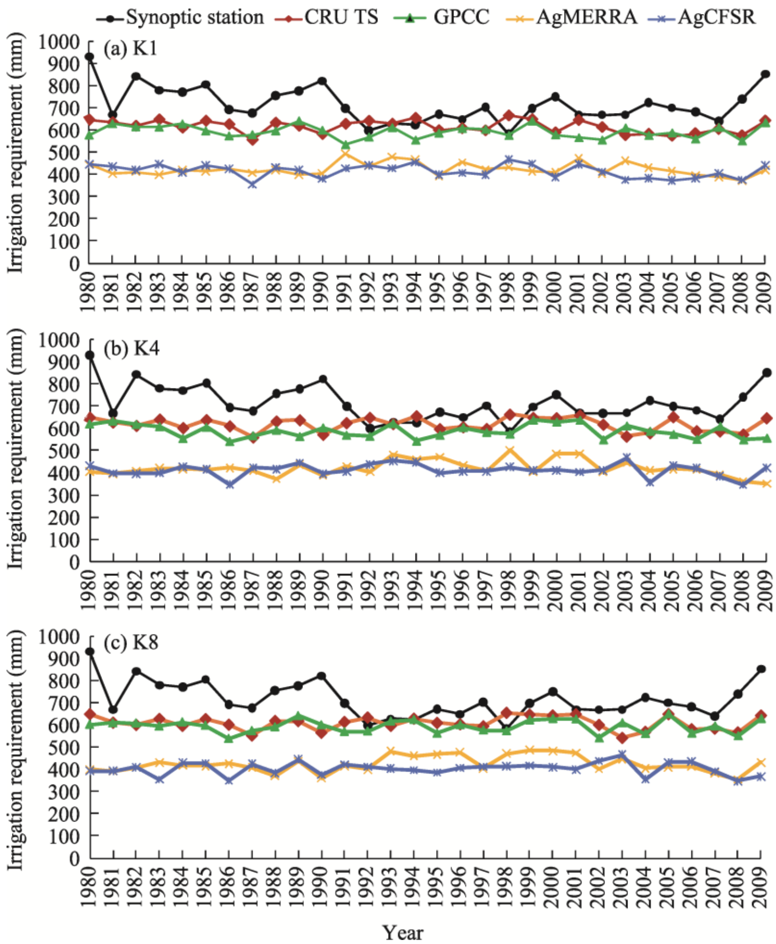

Fig. 6 Comparison of irrigation requirement estimated by CRU TS, GPCC, AgMERRA, and AgCFSR using K1 (a), K4 (b), and K8 (c) with the data from the synoptic station in Qazvin Province during 1980–2009

3.2.3 Biomass

Biomass estimation of each dataset used in this study and the data from synoptic station in Qazvin Province were represented in Figure 7. According to this figure, the results differed significantly from the previous parts (irrigation requirement and reference evapotranspiration). The results of statistical indices were also available in Table 4. According to this table, all datasets were close to each other with a good RMSE. The values of NRMSE for all datasets were less than 5.00%, with an average of about 2.50%. We cannot choose the best dataset here as all of them were close and differ based on K1, K4, and K8, but GPCC using K8 was responsible for the best technique for estimating maize biomass.

3.2.4 Yield

To estimate the yield of maize in this study, it was assumed that irrigation requirement was fully met and the crop was not under any water and salinity stress. The results of each dataset and the

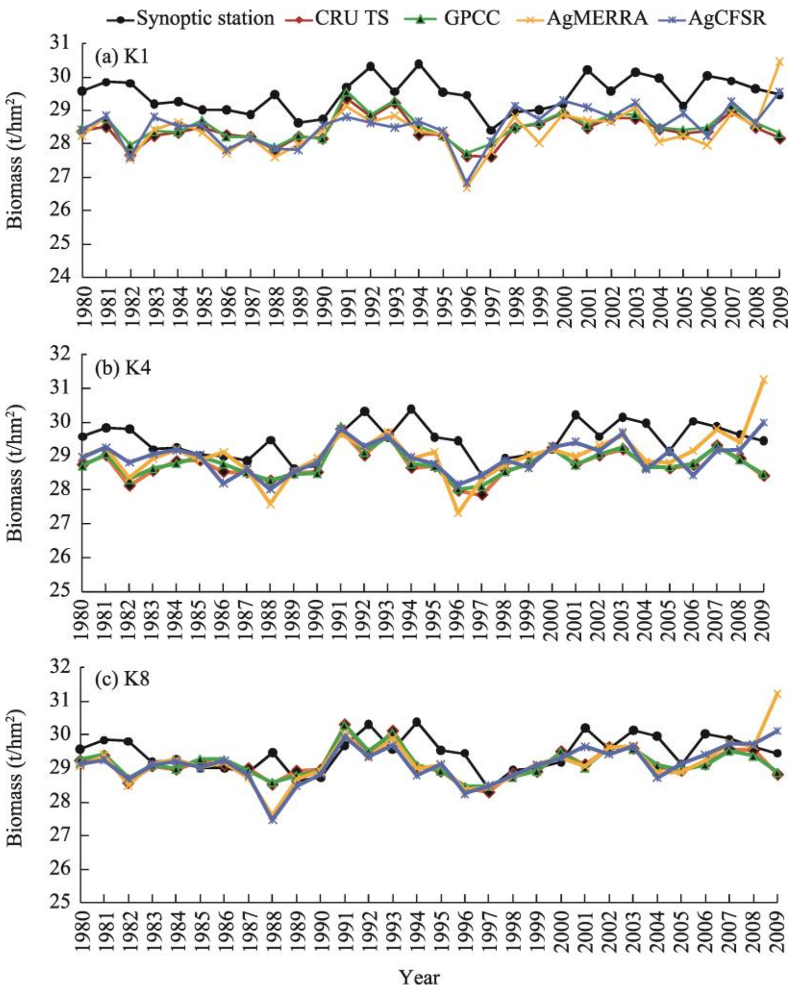

Fig. 7 Comparison of biomass estimated by CRU TS, GPCC, AgMERRA, and AgCFSR using K1 (a), K4 (b), and K8 (c) with the data from the synoptic station in Qazvin Province during1980–2009

synoptic station for estimating the yield of maize were demonstrated (Fig. 8). As can be seen in this figure, the estimations were close to each other. According to Table 4, which provided information about the statistical results of indices, the value of RMSE was close to each other, NRMSE was also less than 5.00%, and the best results go to GPCC using K8.

In short, it can be concluded that all datasets efficiently estimated the biomass and yield of maize in Qazvin Province. For reference evapotranspiration and irrigation requirements, CRU TS was the most efficient dataset. In addition, increasing the number of measurements (i.e., K4 and K8) increased the reliability of the results.

4 Discussion

Bahrololoum et al. (2020) used CRU TS, AgMERRA, AgCFSR, and GPCC to estimate yield and irrigation requirement of wheat in Qazvin Province and to compare the results with the data from the synoptic station in the province from 1980 to 2009 with the same three methods of this study (i.e., K1, K4, and K8). They showed that R^2 and RMSE of CRU TS for estimating evapotranspiration were 0.35 and 87.00 mm for K1, respectively, 0.25 and 91.00 mm for K4, respectively, and 0.27 and 201.00 mm for K8, respectively. Results were not as efficient as CRU TS for AgMERRA, which were less than 0.20 for R^2 and more than 125.00 mm for RMSE for all the three methods. The results of AgCFSR were relatively close to AgMERRA. But GPCC was

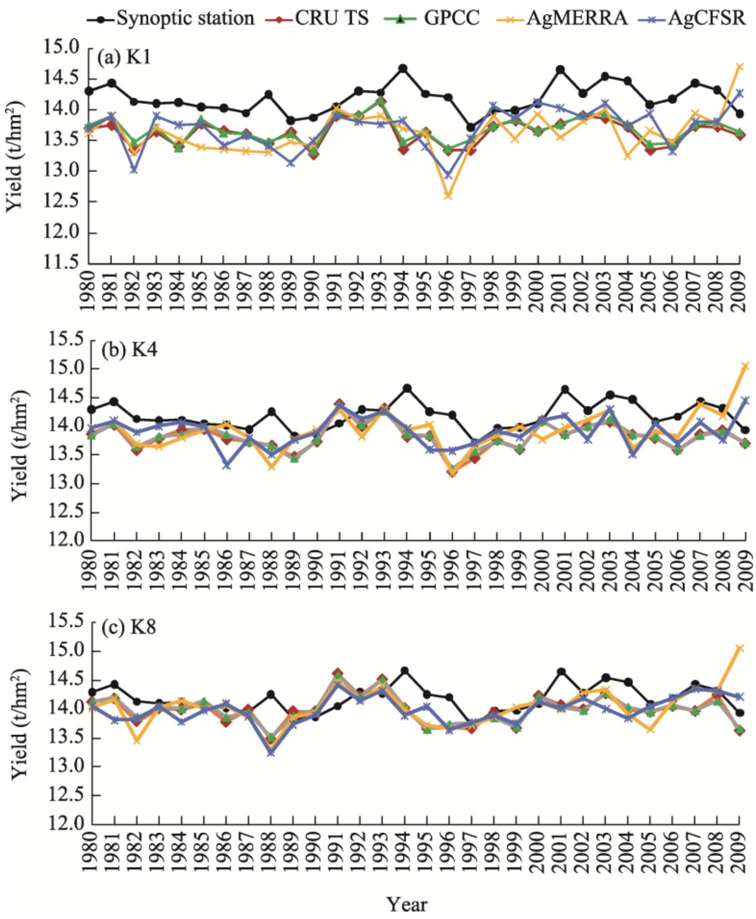

Fig. 8 Comparison of yield estimated by CRU TS, GPCC, AgMERRA, and AgCFSR using K1 (a), K4 (b), and K8 (c) with the data from the synoptic station in Qazvin Province during 1980–2009

the most efficient for estimating the evapotranspiration of wheat. The results were not similar to estimate wheat yield, and AgCFSR was the most efficient among the four datasets with RMSE around 0.22 t/hm^2 and R^2 above 0.70 for K1, K4, and K8. CRU TS was also efficient, but GPCC and AgMERRA were not as efficient as CRU TS and AgCFSR.

Ramezani Etedali and Ahmadi (2021) used Global Land Data Assimilation System (GLDAS), GLDAS-CRU, GLDAS-AgMERRA, and AgCFSR networked meteorological datasets to estimate the yield and evapotranspiration of two crops with a comparison of the results to synoptic station in Qazvin Province from 1980 to 2010 and showed that AgMERRA and AgCFSR were not as efficient as CRU TS and GLDAS for estimating the potential evapotranspiration.

In another study, Gorgin Paveh et al. (2020) showed that CRU TS was more efficient than AgMERRA for estimating wheat's blue and green water footprint. Kakvand et al. (2020) found that GPCC was more reliable than AgCFSR for estimating maize's green and blue water footprint. They also concluded that both datasets were more efficient for estimating the green water footprint (Kakvand et al., 2020). In a recent study, Ramezani Etedali et al. (2022) showed that GPCC and AgMERRA were used to estimate the green, blue, and the total water footprint of maize in Qazvin Province and showed that GPCC was more reliable than AgMERRA. They also showed higher reliability to estimate maize's green and total water footprint than the blue one.

Lashkari et al. (2018), in a study assessing the quality of AgMERRA, found that AgMERRA had a strong relationship with maximum temperature in a province of Iran. Four meteorological datasets (AgMERRA, Asian Precipitation-Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE), Precipitation Estimation from Remotely

Sensed Information using Artificial Neural Networks (PERSIANN) and TRMM) were used in a study related to the estimation of precipitation over monsoon Asia. According to the results, Ceglar et al. (2017) found that AgMERRA was a reliable model based on its structures and its high level of statistical post-processing.

Ahmed et al. (2019) reported that GPCC, CRU, APHRODITE, and University Delaware (UDel) datasets were evaluated in arid regions of Pakistan. They showed that these datasets differed in different climatic regions; however, GPCC performed much better in all climatic regions. They indicated that the use of a significantly larger number of observed stations in the creation of these gauge-based gridded precipitation data was one of the critical reasons for the higher accuracy of GPCC. Similar results were reported by another research related to the number of gauge-based gridded precipitation products using conventional statistical methods (Duethmann et al., 2013).

Finally, it should be mentioned that it is well understood that terrain, wind speed, and hill aspect all significantly impact precipitation. As a result, the gridded precipitation might incorrectly calculate the precipitation distribution (Daly et al., 1994; Bosilovich et al., 2008). Also, elevation differences may significantly impact precipitation and have been extensively researched in previous studies (Johansson and Chen, 2003; Valéry et al., 2010; Yang et al., 2014). For this reason, White et al. (2008) suggested further studies on seasonal patterns of precipitation regionally.

5 Conclusions

According to the importance of accurate estimation of climate variables and crop variables as well as the need to use satellite data instead of classical measurements, in this study, we evaluated four meteorological datasets, i.e., CRU TS, GPCC, AgMERRA, AgCFSR, for estimating climate variables (minimum temperature, maximum temperature, and precipitation) and crop variables (reference evapotranspiration, irrigation requirement, biomass, and yield) in Qazvin Province from 1980 to 2009. For estimating climate variables, the results showed that all datasets were suitable enough to replace the synoptic station in this province. CRU TS was the best dataset for estimating minimum temperature. For estimating maximum temperature and precipitation, AgMERRA and AgCFSR showed more efficiency. For estimating crop variables, the results showed that all datasets were efficient for estimating the yield and biomass of maize. The results were not as efficient for reference evapotranspiration and irrigation requirement, but still can be reliable for CRU TS and GPCC. In addition, the more points we measured from the datasets, the more accurate the results can be obtained. According to the results of this study, we can use meteorological datasets for estimating the climate and crop variables. Furthermore, we can use these datasets in managerial conditions, such as water resources management and agriculture management. Moreover, these datasets can be used to monitor past variations and future trends. It is strongly recommended that more datasets should be investigated in different regions with different climatic conditions and different crops.

Acknowledgements

The authors would like to express their deep gratitude to FAO for providing the AquaCrop model. They also thank the freely available data from the four datasets used in this study, including CRU TS, GPCC, AgMERRA, and AgCFSR.

References

Abedinpour M, Sarangi A, Rajput T B S, et al. 2012. Performance evaluation of AquaCrop model for maize crop in a semi-arid environment. Agricultural Water Management, 110: 55–66.

Ahmed K, Shahid S, Wang X, et al. 2019. Evaluation of gridded precipitation datasets over arid regions of Pakistan. Water, 11(2): 210, doi: 10.3390/w11020210.

Allen R G, Pereira L S, Raes D, et al. 1998. Crop evapotranspiration-Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. Rome: FAO, 300(9): D05109.

Apan A A, Raine S R, Le Brocque A, et al. 2004. Spatial prioritization of revegetation sites for dryland salinity management: An analytical framework using GIS. Journal of Environmental Planning and Management, 47(6): 811–825.

Bahrololoum R, Ramezani Etedali H, Azizian A, et al. 2020. Use of gridded weather datasets in simulation of wheat yield and water requirement (Case study: Iran's Qazvin Plain). Iranian Journal of Ecohydrology, 7(3): 691–706.

Benevides P, Catalao J, Nico G. 2019. Neural network approach to forecast hourly intense rainfall using GNSS precipitable water vapor and meteorological sensors. Remote Sensing, 11(8): 966, doi: 10.3390/rs11080966.

Bosilovich M G, Chen J, Robertson F R, et al. 2008. Evaluation of global precipitation in reanalyses. Journal of Applied Meteorology and Climatology, 47(9): 2279–2299.

Bowen G J, Cai Z, Fiorella R P, et al. 2019. Isotopes in the water cycle: Regional- to global-scale patterns and applications. Annual Review of Earth and Planetary Sciences, 47: 453–479.

Ceglar A, Toreti A, Balsamo G, et al. 2017. Precipitation over monsoon Asia: A comparison of reanalyses and observations. Journal of Climate, 30(2): 465–476.

Confalonieri R, Orlando F, Paleari L, et al. 2016. Uncertainty in crop model predictions: what is the role of users? Environmental Modelling & Software, 81: 165–173.

Daly C, Neilson R P, Phillips D L. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. Journal of Applied Meteorology and Climatology, 33(2): 140–158.

Dorenbos J, Kassam A H. 1979. Yield Response to Water. In: Food and Agriculture Organization of the United Nations. Rome, Italy.

Duethmann D, Zimmer J, Gafurov A, et al. 2013. Evaluation of areal precipitation estimates based on downscaled reanalysis and station data by hydrological modelling. Hydrology and Earth System Sciences, 17(7): 2415–2434.

El-mageed A, Ibrahim M M, Elbeltagi A M. 2017. The effect of water stress on nitrogen status as well as water use efficiency of potato crop under drip irrigation system. Misr Journal of Agricultural Engineering, 34(3): 1351–1374.

Ge Y, Zhang K, Yang X. 2019. A 110-year pollen record of land use and land cover changes in an anthropogenic watershed landscape, eastern China: Understanding past human-environment interactions. Science of the Total Environment, 650: 2906–2918.

Golabi M, Naseri A A. 2015. Assessment Aquacrop model to predict the sugarcane yield and soil salinity profiles under salinity stress. Iranian Journal of Soil and Water Research, 4(46): 685–694.

Gorgin Paveh F, Ramezani Etedali R, Kakvand P. 2020. Estimation of Wheat Water Footprint Based on CRU and AgMERRA Gridded Datasets. [2022-06-23]. https://osau.edu.ua/en/iv-mizhnarodnyj-yevrazijskyj-kongres-z-silskogo-gospodarstva-ta-pryrodnychyh-nauk/.

Harris I, Jones P D, Osborn T J, et al. 2014. Updated high-resolution grids of monthly climatic observations–the CRU TS3.10 Dataset. International Journal of Climatology, 34(3): 623–642.

Harris I, Osborn T J, Jones P, et al. 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Scientific Data, 7(1): 1–18.

Hellal F, Mansour H, Abdel-Hady M, et al. 2019. Assessment water productivity of barley varieties under water stress by AquaCrop model. AIMS Agriculture and Food, 4(3): 501–517.

Javanmard S, Yatagai A, Nodzu M I, et al. 2010. Comparing high-resolution gridded precipitation data with satellite rainfall estimates of TRMM-3B42 over Iran. Advances in Geosciences, 25: 119–125.

Johansson B, Chen D. 2003. The influence of wind and topography on precipitation distribution in Sweden: Statistical analysis and modelling. International Journal of Climatology: A Journal of the Royal Meteorological Society, 23(12): 1523–1535.

Kakvand P, Ramezani Etedali R, Gorgin Paveh F. 2020. Estimation of Maize Water Footprint Based on GPCC and AgCFSR Gridded Datasets. [2022-06-23]. https://osau.edu.ua/en/iv-mizhnarodnyj-yevrazijskyj-kongres-z-silskogo-gospodarstva-ta-pryrodnychyh-nauk/.

Katiraie-Boroujerdy P S, Nasrollahi N, Hsu K. et al. 2013. Evaluation of satellite-based precipitation estimation over Iran. Journal of Arid Environments, 97: 205–219.

Keating B A, Carberry P S, Hammer G L, et al. 2003. An overview of APSIM, a model designed for farming systems simulation. European Journal of Agronomy, 18(3–4): 267–288.

Kreuzer D, Munz M, Schlüter S. 2020. Short-term temperature forecasts using a convolutional neural network–An application to different weather stations in Germany. Machine Learning with Applications, 2: 100007, doi: 10.1016/j.mlwa.2020.100007.

Lashkari A, Salehnia N, Asadi S, et al. 2018. Evaluation of different gridded rainfall datasets for rainfed wheat yield prediction in an arid environment. International Journal of Biometeorology, 62(8): 1543–1556.

Li J, Tian L, Wang Y, et al. 2021. Optimal sampling strategy of water quality monitoring at high dynamic lakes: A remote sensing and spatial simulated annealing integrated approach. Science of the Total Environment, 777: 146113, doi: 10.1016/j.scitotenv.2021.146113.

Massah A R, Morid S. 2005. Effects of climate change on Zayandeh Rud river flows. Journal of Science and Technology of Agriculture and Natural Resources, 4: 17–27.

Mitchell T D, Carter T R, Jones P D, et al. 2004. A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100). Geography, 55: 25, doi: 10.1002/joc.1181.

Mitchell T D, Jones P D. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology, 25(6): 693–712.

New M, Hulme M, Jones P. 1999. Representing twentieth-century space-time climate variability. Part I: Development of a 1961-90 mean monthly terrestrial climatology. Journal of Climate, 12(3): 829–856.

New M, Hulme M, Jones P. 2000. Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate. Journal of Climate, 13(13): 2217–2238.

New M, Todd M, Hulme M, et al. 2001. Precipitation measurements and trends in the twentieth century. International Journal of Climatology, 21(15): 1889–1922.

Olivera Rodriguez P, Holzman M E, Degano M F, et al. 2021. Spatial variability of the green water footprint using a medium-resolution remote sensing technique: The case of soybean production in the Southeast Argentine Pampas. Science of the Total Environment, 763: 142963, doi: 10.1016/j.scitotenv.2020.142963.

Ollivier C, Olioso A, Carrière S D, et al. 2021. An evapotranspiration model driven by remote sensing data for assessing groundwater resource in karst watershed. Science of the Total Environment, 781: 146706, doi: 10.1016/j.scitotenv.2021. 146706.

Ordikhani H, Parashkoohi M G, Zamani D M, et al. 2021. Energy-environmental life cycle assessment and cumulative exergy demand analysis for horticultural crops (Case study: Qazvin province). Energy Reports, 7: 2899–2915.

Queyrel W, Habets F, Blanchoud H, et al. 2016. Pesticide fate modeling in soils with the crop model STICS: Feasibility for assessment of agricultural practices. Science of the Total Environment, 542: 787–802.

Ramezani Etedali H, Ahmadi M. 2021. Evaluation of various meteorological datasets in estimation yield and actual evapotranspiration of wheat and maize (case study: Qazvin plain). Agricultural Water Management, 256: 107080, doi: 10.1016/j.agwat.2021.107080.

Ramezani Etedali H, Gorgin F, Kakvand P. 2022. Study of the performance of two meteorological datasets in estimating the maize water footprint, a case study: Qazvin Plain. Iranian Journal of Irrigation and Drainage, 15(6): 1394–1403.

Reichle R H, Koster R D, De Lannoy G J M, et al. 2011. Assessment and enhancement of MERRA land surface hydrology estimates. Journal of Climate, 24(24): 6322–6338.

Rodriguez J, Ustin S, Sandoval-Solis S, et al. 2015. Food, water, and fault lines: Remote sensing opportunities for earthquake-response management of agricultural water. Science of the Total Environment, 565: 1020–1027.

Ruane A C, Goldberg R, Chryssanthacopoulos J. 2015. Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. Agricultural and Forest Meteorology, 200: 233–248.

Saeidi R, Ramezani Etedali H, Sotoodehnia A, et al. 2021. Evaluation of AquaCrop model for estimating of changes process of soil moisture, evapotranspiration and yield of maize under salinity and fertility stresses. Environmental Stresses in Crop Sciences, 14(1): 195–210.

Salehnia N, Zare H, Kolsoumi S, et al. 2018. Predictive value of Keetch-Byram Drought Index for cereal yields in a semi-arid environment. Theoretical and Applied Climatology, 134(3): 1005–1014.

Schneider U, Becker A, Finger P, et al. 2014. GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. Theoretical and Applied Climatology, 115(1): 15–40.

Shi H, Li T, Wei J. 2017. Evaluation of the gridded CRU TS precipitation dataset with the point raingauge records over the Three-River Headwaters Region. Journal of Hydrology, 548: 322–332.

Stöckle C O, Donatelli M, Nelson R. 2003. CropSyst, a cropping systems simulation model. European Journal of Agronomy, 18(3–4): 289–307.

Strangeways I. 2006. Precipitation: Theory, Measurement and Distribution. Cambridge: Cambridge University Press.

Valéry A, Andréassian V, Perrin C. 2010. Regionalization of precipitation and air temperature over high-altitude catchments–learning from outliers. Hydrological Sciences Journal–Journal des Sciences Hydrologiques, 55(6): 928–940.

Vorosmarty C J, Sahagian D. 2000. Anthropogenic disturbance of the terrestrial water cycle. BioScience, 50(9): 753–765.

White J L, Knapp A K, Kelly E F. 2008. Increasing precipitation event size increases aboveground net primary productivity in a semi-arid grassland. Oecologia, 158(1): 129–140.

WMO, ICSU. 1990. The Global Precipitation Climatology Project–Implementation and data management plan. [2022-08-17]. https://library.wmo.int/index.php?lvl=notice_display&id=11757#.Y4ZbLnZBzIU.

Worqlul A W, Collick A S, Tilahun S A, et al. 2015. Comparing TRMM 3B42, CFSR and ground-based rainfall estimates as input for hydrological models, in data scarce regions: the Upper Blue Nile Basin, Ethiopia. Hydrology and Earth System Sciences Discussions, 12(2): 2081–2112.

Yang H S, Dobermann A, Lindquist J L, et al. 2004. Hybrid-maize–A maize simulation model that combines two crop modeling approaches. Field Crops Research, 87(2–3): 131–154.

Yao T, Xue Y, Chen D, et al. 2019. Recent third pole's rapid warming accompanies cryospheric melt and water cycle intensification and interactions between monsoon and environment: Multidisciplinary approach with observations, modeling, and analysis. Bulletin of the American Meteorological Society, 100(3): 423–444.

Zhu X, Xu K, Liu Y, et al. 2021. Assessing the vulnerability and risk of maize to drought in China based on the AquaCrop model. Agricultural Systems, 189: 103040, doi: 10.1016/j.agsy.2020.103040.